

SEARCH BASED MASHUP/DASHBOARDS: ARCHITECTURE AND BEST PRACTICE

Author

Eduard Daoud
interface projects GmbH
Zw inglistraÙe 11/13
01277 Dresden
Germany
eduard.daoud@interface-projects.de

1. Abstract.....	1
2. Context.....	1
3. What is Mashup?.....	3
4. What is Search based Mashup?.....	3
5. How SBMs Are Being Used.....	3
6. Challenges for search based mashup	3
7. Semantic Descriptions in SBM.....	4
8. SBM Architecture.....	5
9. Lessons learned and best practices.....	6
10. Conclusions.....	7
11. Acknowledgement.....	7
12. Bibliography.....	7

1. Abstract

The goal of this white paper is to provide insight into one of the most rapidly de-veloping sectors in today's business. The paper discusses the architecture and use in practice of the search based mashup, as well as the benefits and advantages compared with other mashup models. Also covered are lessons learned and best practices from the development of a search based, metadata-focused information retrieval mashup called intergator dashboard. In particular, we discuss the chal-enges and possible solutions in an enterprise search scenario, and show where semantic descriptions and information access matter in such a solution.

Keywords: search based mashup, search based applications, information access, Dashboards, InfoApps.

2. Context

Information management is becoming increasingly important and is the fastest growing sector in today's business. It has a special role within an organisation when it comes to making information and data from the various systems available for employees and management. On the one hand we talk about big data issues, but in reality the knowledge within a company and the knowledge of its employees is the goal. As one of these rapid developments, knowledge management as part of holistic information management strategies is focused on search-based approaches. [1] The development of editorial software components that query the explicit knowledge of the employees and security in the data structures will be made available with the introduction of an enterprise search platform (*enterprise search*) that aggregates and centralizes all of the data from various sources. Once the data is aggregated, the software evaluates, analyses and understands data from various structures and applications. Two of the biggest challenges are the types of data (structured, unstructured and semi-structured) and the expectations of users. A key to the solution for this dilemma lies in the enhancement of the intranet for centralized, interactive and personalized search based information mashup. From a technical point of view, it is about semantic mashups in the form of InfoApps (search index-based applications). Lightweight InfoApps assume the visualization, access and personalization of content. We present the approach in this paper as follows: After describing related works, the proposed mashup approach and the challenges for search based mashups (SBM) in Sections **Fehler! Verweisquelle konnte nicht gefunden werden.**, **Fehler! Verweisquelle konnte nicht gefunden werden.** & **Fehler! Verweisquelle konnte nicht gefunden werden.**, we discuss the

semantic descriptions and machine learning in SBM. In Sections **Fehler! Verweisquelle konnte nicht gefunden werden.** & **Fehler! Verweisquelle konnte nicht gefunden werden.** we discuss the maturity of the framework *intergator* and how to interact with the dashboard & mashups. In Sections **Fehler! Verweisquelle konnte nicht gefunden werden.** & **Fehler! Verweisquelle konnte nicht gefunden werden.** we provide an overview of the lessons learned and best practices, and also discuss the possibilities for demonstration of the software. Finally, in Section **Fehler! Verweisquelle konnte nicht gefunden werden.** we offer some concluding remarks and outline future works.

3. What is Mashup?

Mashup tools are interactive application and framework which focused and target the needs of end-users developing a specific kind of Web applications, built of the reuse and composition of multiple (web) data sources and Web services [2]. In the last 5 years have rapid growth many mashup tools especially in industry offering a range of information access capability.

Some of the mashup tools are based on visual component, others feature focused on automation business process support, many of mashups support collaborative development engaging public and private online communities. [2].

A revolutionary model is which reduces IT costs and complexity by using Information Access based on powerful enterprise search. Firstly we should deep dive in term search based solution to understand this approach.

4. What is Search based Mashup?

A search-based Mashup (SBM) is InfoApps running in a dashboard built upon a platform that's capable of decoupling data with connectors from its original source. Unlike traditional software applications that can only query content in a database, an SBM can query the index which building from disparate structured, semi-structured and unstructured content sources and return the results of the query in a single, unified view. [4]

Software applications in which a search engine platform is used as the core infrastructure for information access and reporting. SBMs use semantic technologies to aggregate, normalize and classify unstructured, semi-structured and/or structured content across multiple repositories, and employ natural language technologies for accessing the aggregated information. Some software vendors use the term SBA-Search based Application. Like EXALEAD¹ and interface projects GmbH.²

A search-based application is generally taken to mean an application that uses a search engine index at its core, rather than a database or other structure. Search engines are extremely good at slicing and dicing information on-the-fly. Search-based applications exploit this capability.

5. How SBMs Are Being Used

SBMs are useful and relevant for all software in which information access (indexing, processing, retrieval, search and present content present) play the core objective.

They are being used to create innovative tailored search applications for the support of business processes within consolidated access to structured and unstructured information search applications can consider the context of customized commercial processes and deliver optimum results and improve database applications. Monitoring of information across organisation, support Intranet with an universal search function. Customized search applications could support also e.g. in quality management, sales, customer service, human resource and Content Analytics. [5]

SBMs provide information access faster and far cheaper than relational database querying. Because search engines are uniquely designed for fast information access ('read' operations) by vast numbers of users against massive data volumes, shifting queries from a database to an index can significantly reduce costs (e.g., database offloading) even as one expands access and improves performance.

6. Challenges for search based mashup

The challenges for search based mashup is on the one hand connectivity, semantic and search performance in a huge amount of data and content and on other hand the privacy.

10.000 pages on a web server, we can do all analysis necessary for a rich search solution. But if we want to index a file server with >100.000.000 documents it is very strong to do more with it as full-text indexing (e.g. in such usage scenarios it is not reason-able to perform detailed linguistic and text structure analysis). Here are limitations for semantic notation of metadata. The quality of data and text extraction tools. This is the ground of a universal search solution. No semantic analysis makes sense if the text and meta-data extraction part is not a stable piece of software. It must run 7x24 hours without crashes and memory leaks (horrible content of many PDF-documents is one example for that).

¹ <http://www.3ds.com/products-services/exalead> Accessed 28 05 2016

² www.interface-projects.de Accessed 28 05 2016

Often different content sources require handling of different authentication schemas. Then security access for search results is the next challenge (think about different user management of databases, file servers and PLM).

Last but not least – complexity of the processing infrastructure and semantic framework which consist of crawlers, modeling tools, analyzers, visualization components and semantic integration parts). Additionally there are limitations for linguistic components in mixed-language scenarios like dealing with IT-content.

General play analysing, aggregation and combining of personal data a central role, these are subject to the privacy policy and there are special rules and requirements in dealing with them.

Search based systems are from the perspective of IT security once common IT systems and therefore represents all that must be observed when operating a server or database system. Even for the aggregation of data, there are privacy-friendly security approaches, which are summarized under the term privacy Preserving [6].

7. Semantic Descriptions in SBM

Semantic mashup systems based traditionally on ontology model or on a manual editing information model. Search based approach focuses on a hybrid model, especially in the context of enterprise search. The hybrid model

delivers an intelligent collection of business objects, entities, and classes to be considered for the systems Meta data. These elements collected throughout the indexing and keep update though the query.

The biggest challenge in semantic mashup is the update process of the ontology behind. Every organization is unique and has its subject specific knowledge domain and therefore requires a specific and tailor-made ontology.

In recent years, deep artificial neural networks have won numerous contests in pattern recognition, especially machine learning, a shift has taken place away from handcrafted, rule-based systems towards statistical models that learn from a huge amounts of data. Machine Learning proceedings were an old concept, but in the past was not the appropriate hardware and computing power to train and learn a model from a huge amounts of data. This approach is possible now, thank the powerful computing technologies like GPU-accelerated computing. The relationships between different concepts can be learned instead of manual editing with the help of ontology editor.

However, the machine learning proceedings have a big challenge if we use it in a closed organization with own knowledge domain. But a hybrid approach like used in considered mashup tool *intergator*³ can be entirely learned based on the combination between an information model and machine learning methods.

³ www.intergator.de

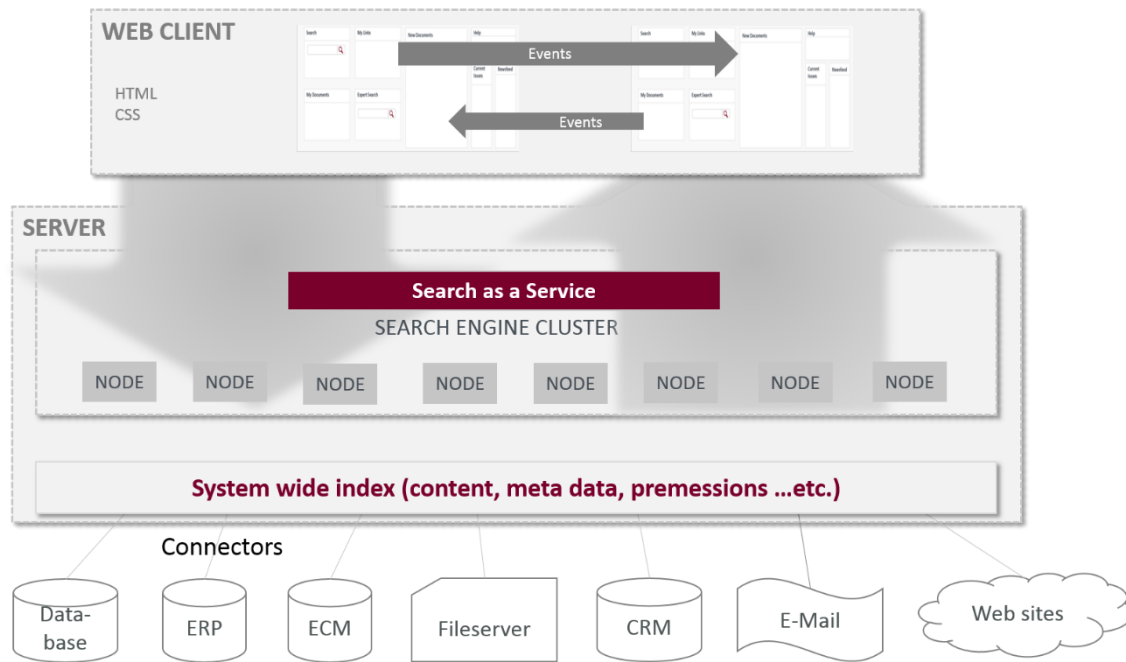


Figure 1 Search based mashup as an Enterprise Data Infrastructure Component

8. SBM Architecture

Back to the considered tool integrator SBM or integrator social Dashboard ⁴. To classify this Tool accordance with the criteria of [7], we should first explain briefly the functioning of the dashboards. The Fundament of the integrator Dashboard is the system wide indexing throughout various systems. Results of preconfigurable search queries is shown in free placeable InfoApps/Widgets on the Dashboard. Dashboard layout is a grid system. InfoApps are placed in columns the length of the App is dynamic the functional focus is the simplest possible personal adaptability of the dashboard so that very individual information needs can be served. See Figure 2 integrator Dashboard

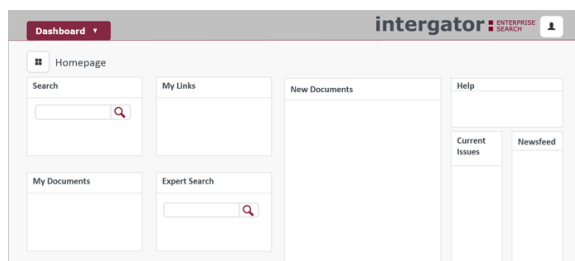


Figure 2 intergator Dashboard

The search-based Architecture assumes integration into the existing infrastructure as

login, SSO and permissions. Now we could classify this Tool accordance with the criteria of [7, p. Chap. 6]. Search based mashups are data mashups with client-side rendering of search results and possible intermediate progression information. But in the same it is possible to place and render content direct from the web or other data source without indexing by using RSS feed or HTML Code as example. See Figure 1 Search based mashup as an Enterprise Data Infrastructure Component

Hereby it does not matter, where these documents have been stored. integrator indexes databases, file servers, Microsoft Sharepoint, Lotus Notes, Wikis, Blogs, DMS, ERPs, CRMs as well as users' email account. By log in via Single Sign-On at the Social Dashboard, all rights of access are considered when you get presented search results. The search-based presentation of business information is completed by particular InfoApps, which can show rss-feeds, inputs on blog-posts and wikipages in real-time or contain links as navigation to relevant areas of the intranet. The choice of apps template can be administered centrally.

⁴ <https://www.intergator.de/en/product/intergator-enterprise-search/social-dashboard/> Accessed 28 05 2016

9. Lessons learned and best practices

A search based mashup is an enterprise mashup⁵ because the integration of heterogeneous digital data and applications from multiple sources for business purposes is the goal.

At the center of the SMB project are the user and his productivity. Therefore the visualization and user interaction are very important.

Search based mashup as Intranet and single Point of Access:⁶



Figure 3 Search based mashups as Intranet

The Bitburger Brewer Group GmbH, with the brands Bitburger, König Pilsener, Köstritzer, Licher and Wernesgrüner, is a family business in its seventh generation. With approximately 1,800 employees, it is one of the leading brewing groups in Germany. A search engine for everyone. Bitburger using a single search engine for Intranet, Extranet and web portals in the Bitburger Brewery Group. The Intranet build on basis of a mashup technology of the product intergator. Every user is able to create his own dashboards and use already preconfigured mashups like SAP-Accounting, Links, news, How-To's etc. This information is stored in different applications across more than 6 locations. The project team is able to create

and share quickly, independent and with support from the IT department and

Comprehensive application based on the search based Mashups SMB.

Search based mashup as compliance Cockpit

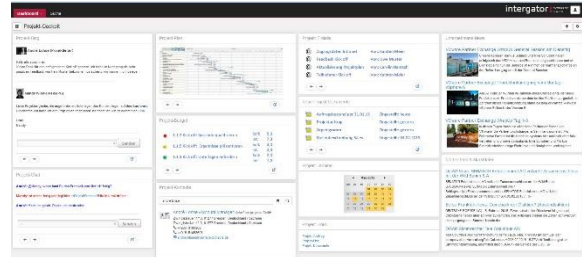


Figure 4 intergator compliance cockpit

The Viessmann Group is one of the world's leading manufacturers of heating, industrial and refrigeration systems. Internal users with help of a search based mashup platform to collect and combine information from different data sources and formats. Noncompliant actions or transactions can be identified by matching all indexed data with defined compliance rules. The result is prepared for the compliance officer and pushing on the cockpit within the mashups. This ensures a comprehensive compliance monitoring.

With help of the concept of single point of access the system provides every employee a compliance dashboard and thus a InfoApps/mashups available in which it checks its data permissions and even corrects errors. No matter where the data is stored, whether that is in a DMS, an e-mail or on the file server, the compliance dashboard keeps track of its relevant data.⁷

⁵ An enterprise mashup is also sometimes known as a business mashup or, less precisely, as a data mashup. [7]

⁶ <https://www.intergator.de/en/wp-content/uploads/sites/2/2015/11/intergator-case-study-bitburger-brewery.pdf> Accessed 28 05 2016

⁷ <https://www.intergator.de/wp-content/uploads/2016/03/intergator-Fallstudie-Appliance-Viessmann.pdf> Accessed 28 05 2016

10. Conclusions

Search based mashup for business or enterprise search based mashup is a robust technology. Example of a SBM is intergator Dashboard. The vender implement a semantic search engine in the core. The retrieved results are also merged and linked each other in the index. Therefore, the different types of contents can be referred in the UI with help of InfoApps. The implemented search engine is trusted from more than 200 big enterprises over the world. The results show usefulness and effectiveness content searches with a variety of content types. This revolutionary model is which reduces IT costs and complexity by using Information Access based on powerful enterprise search is one of the future promising implementations of mashups.

11. Acknowledgement

The work presented in this paper has been supported by the interface projects GmbH. The author would like to thank Mr. Dr. Uwe Crenze, for his effort to share the success stories to conduct experiments.

12. Bibliography

- [1] Transparency Market Research (TMR), "Enterprise Search Market (Government and commercial offices, Banking and financial services, Retail, Healthcare and Others) - Global Industry Analysis, Size, Share, Growth, Trends and Forecast," Albany NY - 12207, United States, 2015.
- [2] D. Benslimane , A. Sheth and S. Dustdar, "Services mashups: The new generation of web applications," 2008.
- [3] R. Fox, J. Cooley and M. Hauswirth, "Collaborative development of trusted mashups," 2010.
- [4] IDC, Analyze the Future, "IDC on The High Cost Of Not Finding Information," IDC, Framingham, MA 01701, 2014.
- [5] interface projects GmbH, "Dashboard datasheet intergator 5," interface projects GmbH, 10 12 2015. [Online]. Available: www.intergtor.de. [Accessed 28 05 2016].
- [6] C. C. Aggarwal and P. S. Yu, Privacy-Preserving Data Mining, New York: Springer Verlag, 2008.
- [7] M. Matera and F. Daniel, Mashups: Concepts, Models and Architectures, Springer, 2014.